AL-TP-1992-0037

AD-A256 413

ARMSTRONG LABORATORY

(2)

THE CORRELATION OF COGNITIVE AND PSYCHOMOTOR TESTS

Malcolm James Ree
Thomas R. Carretta

HUMAN RESOURCES DIRECTORATE
MANPOWER AND PERSONNEL RESEARCH DIVISION
Brooks Air Force Base, TX 78235-5000

September 1992

Interim Technical Paper for Period January 1991 – June 1991

Approved for public release; distribution is unlimited.

AIR FORCE MATERIEL COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235-5000

NOTICES

This paper is published as received and has not been edited by the technical editing staff of the Armstrong Laboratory.

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Office of Public Affairs has reviewed this paper, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This paper has been reviewed and is approved for publication.

MALCOLM JAMES REE
Project Scientist

WILLIAM E. ALLEY, Technical Director
Manpower and Personnel Research Division

ROGER W. ALFORD, Lt Col, USAF
Chief, Manpower and Personnel Research Division

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 1992	Interim – January 1991 – June 1991

4. TITLE AND SUBTITLE
The Correlation of Cognitive and Psychomotor Scores

5. FUNDING NUMBERS
PE – 62205F
PR – 7719
TA – 18
WU – 45

6. AUTHOR(S)
Malcolm James Ree
Thomas R. Carretta

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Armstrong Laboratory
Human Resources Directorate
Manpower and Personnel Research Division
Brooks Air Force Base, TX 78235-5000

8. PERFORMING ORGANIZATION REPORT NUMBER
AL-TP-1992-0037

9. SPONSORING/MONITORING AGENCY NAMES(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

A paper-and-pencil multiple-aptitude test battery and a computer-based psychomotor test battery were administered to a sample of 354 Air Force recruits. The tests of the multiple-aptitude battery were used to estimate psychometric *g* and to predict the psychomotor tests. The multiple correlation of the multiple-aptitude tests and each psychomotor test as a criterion was .50, corrected for range restriction. The average correlation of the psychomotor tests and psychometric *g*, corrected for range restriction and unreliability, was .73. The multiple-aptitude tests and the psychomotor tests were correlated and subjected to a principal components analysis. The average *g* saturations of the psychomotor and multiple-aptitude tests were .76 and .87, respectively. Confirmatory factor analyses disclosed hierarchical general cognitive and general psychomotor factors, two lower order multiple-aptitude test factors and three lower order psychomotor test factors.

14. SUBJECT TERMS
Factor structure
Psychometric *g*
Psychomotor ability

15. NUMBER OF PAGES
28

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

CONTENTS

FIGURES

TABLES

PREFACE

This research and development was accomplished under work unit 77191845 which is dedicated to the study of selection and classification of Air Force Pilots. The authors thank Laurie C. Walters, Melody Darby, and Rick Siem for help in planning the data collection and in statistical analyses. We thank James Earles and Major David Perry for reading drafts and 1Lt Laura Kreitler for producing the figure of the structural model. Comments from Joseph Weeks on the correction for range restriction which helped clarify our thinking were appreciated. Professor Charles Lance, University of Georgia, is thanked for his helpful comments on statistical tests appropriate for the structural models. A special debt of gratitude is owed William C. Tirre for his generosity of time in developing the structural model.

"Facts are very stubborn things, overruling all theories."
Jules Verne, *Journey to the Center of the Earth.*

THE CORRELATION OF COGNITIVE AND PSYCHOMOTOR TESTS

SUMMARY

A study was conducted to investigate the nexus of cognitive and psychomotor tests which are frequently seen as taxonomically independent. A paper-and-pencil multiple-aptitude test battery and a computer based psychomotor test battery were administered to a sample of 354 Air Force recruits. The tests of the multiple-aptitude battery were used to estimate psychometric *g* and to predict the psychomotor tests. Cognitive multiple-aptitude tests were found to be related to psychomotor scores. The multiple correlation of the cognitive tests and each psychomotor test as a criterion was .50, corrected-for-range-restriction. The average correlation of the psychomotor tests and psychometric *g*, corrected-for-range-restriction and unreliability, was .73. The cognitive tests and the psychomotor tests were correlated and subjected to a principal components analysis. The average *g* saturations of the psychomotor and cognitive tests were .76 and .87 respectively. Confirmatory factor analyses disclosed hierarchical general cognitive and general psychomotor factors, two lower order cognitive test factors and three lower order psychomotor factors. The most important practical implication is that the higher order psychomotor factor needs to be studied to determine its role in predictive validity. This could shape future development of applied psychomotor tests.

INTRODUCTION

Most psychologists view cognitive and psychomotor abilities as two distinct and independent categories (Fleishman & Quaintance, 1984, p. 162) although within each category there may be several factors. Multiple-aptitude batteries, cognitive measures, are frequently factored and among the factors reported are psychometric *g*, verbal, quantitative, spatial, perceptual speed, and technical information (Jensen, 1980; Kass, Mitchell, Grafton & Wing, 1983; Ree & Earles, 1991a; Skinner & Ree, 1987). Factor analyses of computer administered elementary cognitive tasks (ECTs) showed factors for psychometric *g*, working memory, reasoning, and reaction time (Carroll, 1991; Kranzler & Jensen, 1991; Kyllonen & Christal, 1990). Psychomotor batteries have yielded factors including control precision, multilimb coordination, reaction time, and rate control (Fleishman, 1953, 1964, 1966) but no higher order factor like psychometric *g*. Cronbach (1970) has stated that, unlike cognitive tests, there is no general psychomotor factor. The dissimilarity of factor names from the different areas probably represents different theoretical and taxonomic perspectives. It is possible that the same constructs are being referred to by different names.

Several studies have shown both paper-and-pencil multiple-aptitude batteries and batteries of ECTs administered by computer to be substantially saturated with a common factor, psychometric *g*. For example, the Armed Services Vocational Aptitude Battery (ASVAB) has been shown to be heavily *g* saturated (Ree & Earles, 1991a) as have other similar tests (Jensen, 1980; Earles & Ree, 1992a). Kranzler and Jensen (1991) have shown that a large battery of ECTs predicts scores on Raven's (1966) Advanced

Progressive Matrices, a test frequently acknowledged as a good marker of *g* (Jensen, 1987; Kranzler & Jensen, 1991; Neubauer, 1991). Kranzler and Jensen (1991) have also shown that the ECTs predict scores on the Multidimensional Aptitude Battery (Jackson, 1984), a test similar to the ASVAB, demonstrating the *g* saturation of ECTs.

Hunter (1980) demonstrated some commonality between paper-and-pencil scores on the General Aptitude Test Battery (GATB) and unrefined manual GATB psychomotor scores. However, the GATB psychomotor tests appear substantially different than those being developed and implemented today (Carretta, 1989). These new psychomotor tests allow precise computer measurement and require manipulation of control sticks rather than the simple manual dexterity required by the GATB psychomotor tests. Given a battery of psychomotor tests and Cronbach's (1970) assertion it is appropriate to ask if a higher order general psychomotor factor exists. Based on Hunter's (1980) finding it is appropriate to ask if psychomotor tests measure *g*.

Although studies of the factor structure of cognitive tests (Ree, Mullins, Mathews, & Massey, 1982), ECTs (Carroll, 1991) or psychomotor tests (Fleishman, 1964, 1966) were available, no studies could be found which simultaneously investigated the factor structure of cognitive and computer-based psychomotor tests. Additionally, no investigations of a general psychomotor factor could be found. The purpose of this study was to investigate the factor structure of a typical multiple-aptitude battery and a typical group of computer-based psychomotor tests, estimate the *g* saturation of the psychomotor tests and determine if a general psychomotor factor existed.

METHOD

Subjects

The subjects were 354 Air Force recruits with a median age of 21 years and were mostly white (78%), male (86%) and high school graduate or better (99%). All subjects were selected for Air Force enlistment, in large part, on the basis of their ASVAB scores and educational achievement.

Measures

The ASVAB is a 10 test multiple-aptitude battey used by the U. S. military. It measures psychometric *g* (Ree & Earles, 1991a) and factors found to be valid for predicting a variety of criteria (Earles & Ree, 1992b; Ree & Earles, 1991b; Ree, Earles, & Teachout, 1992). The tests include: General Science (GS), Arithmetic Reasoning (AR), Word Knowledge (WK), Paragraph Comprehension (PC), Numerical Operations (NO), Coding Speed (CS), Auto & Shop Information (AS), Mathematics Knowledge (MK), Mechanical Comprehension (MC), and Electronics Information (EI). The NO and CS tests are speeded, all the others are power. Psychometric *g* was computed as suggested by Ree and Earles (1991a).

The psychomotor tests were from the Basic Attributes Test (BAT), a battery of computer based tests validated for selection of candidates for pilot training (Carretta, 1989, 1990). The BAT was computer administered with a special alpha-numeric key pad, a monochrome monitor and two control (joy) sticks. The first psychomotor test was a rotary pursuit task called Two Hand Coordination, an example of Fleishman's control precision (Fleishman & Quaintance, 1984). In this test the subject used right and left hand control sticks to keep a circle on a representation of an airplane as it moved in an ellipse on the computer monitor. The two scores computed were horizontal tracking distance error (THH) and vertical tracking distance error (THV). Complex Coordination, an example of multilimb coordination (Fleishman & Quaintance, 1984) was the second psychomotor test. Using the right hand control stick, this compensatory tracking task required the subject to keep a 1" cross centered on a dotted-line cross which bisected the monitor horizontally and vertically. Simultaneously, using the left hand control stick, the subject had to keep a 1" vertical bar horizontally centered at the base of the monitor display. The 1" cross and the vertical bar were forced away from center by a random function. The three scores for this test were horizontal tracking distance error (CCH) and vertical tracking distance error (CCV) for the 1" cross and tracking distance error (CCR) for the 1" vertical bar. The third psychomotor test, Time Sharing, was identified with Fleishman & Quaintance's (1984) psychomotor factors of reaction time and rate control. In the first 10 minutes, the subject was required to keep randomly-moving cross-hairs on an airplane target using the right hand control stick. In the next 6 minutes the subject had to repeat the tracking task and had to cancel digits which appeared at random intervals and positions on the monitor. Cancellation was timed and consisted of pressing the corresponding digit on the numeric keypad. Tracking task difficulty was computer adjusted to maintain task load equivalent across subjects. Smaller tracking errors caused the stick sensitivity to increase and larger tracking errors caused it to decrease. The three scores on this test were tracking difficulty on the task without digit cancellation (TSS), digit cancellation reaction time (TSR), and tracking difficulty during digit cancellation (TSD). Electro-mechanical versions of these psychomotor tests were used during the second World War. These computer administered versions are analogues of the electro-mechanical ones reported by Thorndike and Hagen (1959). A detailed description of the BAT was provided by Carretta (1987). Correlations involving error and response time scores were reflected so that good performances were always positively correlated.

Procedure

The ASVAB was administered as part of the operational enlistment qualification procedures and the BAT was administered on the 11th day of basic military training. The subjects were told that the BAT scores were being collected for experimental purposes only and although given the opportunity to decline participation, none did.

Analyses included descriptive statistics, correlations, principal component analyses and confirmatory factor analysis. The matrix of correlations of ASVAB and BAT tests was corrected-for-range-restriction by the method of Lawley (1943) to the normative sample (Maier & Sims, 1986; Ree & Wegner, 1990). Each psychomotor test score was predicted by the set of 10 ASVAB tests as a measure of commonality with cognitive ability. An unrotated principal components analysis (Hotelling, 1933a, 1933b)

was conducted to estimate g loadings (Jensen, 1980; Ree & Earles, 1991a) of all the tests. To determine if the first component were still g as in previous studies (Ree & Earles, 1991a), coefficients of congruence (Burt, 1948) were computed between the loadings for the cognitive tests estimated in the presence of the psychomotor tests and the loadings for the cognitive tests estimated without the psychomotor tests in a different sample. The loadings on the first unrotated principal component were divided by the square roots of the test-retest reliabilities (Carretta, in press; Earles & Ree, in press) of the tests (Jensen, 1980) to better estimate the true g loadings. All multiple and single correlations were tested at $p < .01$ Type I error rate.

Several confirmatory factor analyses were conducted on the matrix of corrected correlations to find a parsimonious model which fit the data well. The first was a simple hierarchical residualized model suggested by the content and administration modes (paper-and-pencil versus computer) of the tests. It consisted of three factors: psychometric g, a general psychomotor factor, and a paper-and-pencil versus computer or administration method factor. A hierarchical residualized orthogonal seven factor model with g, a general psychomotor factor, two paper-and-pencil factors suggested by test content, and three psychomotor factors suggested by Fleishman's factors (Fleishman & Quaintance, 1984) was tried. The final hierarchical residualized orthogonal model had nine factors including g, a general psychomotor factor, four paper-and-pencil factors suggested by Kass, Mitchell, Grafton, & Wing (1983), and three psychomotor factors suggested by Fleishman.

Several fit statistics including the Bentler-Bonett non-normed fit statistic (Bentler, 1989), χ^2/I2 (Marsh, Balla, & McDonald, 1988), the Tucker-Lewis fit index (Tucker & Lewis, 1973) and the standardized residuals were evaluated to determine the best fitting model. Marsh, Balla, and McDonald (1988) have shown that the Bentler-Bonett index may be susceptible to sample size effects. They recommend evaluation of the χ^2/I2 index, which tests the substantive model against a null model with as many factors as variables (the independence model), and the Tucker-Lewis TLI incremental fit index to determine the most appropriate factor structure.

RESULTS

Computation of descriptive statistics of the ASVAB tests showed that the sample was range restricted. The average ASVAB test scores were about one half of a standard deviation above the population mean (Table 1).

No population norms exist for the psychomotor tests. Table 2 shows the correlations among the ASVAB and BAT tests both as observed and corrected-for-range-restriction. Appendix A provides a brief explanation of the multivariate correction for range restriction used in this study.

Table 1. Means and Standard Deviations of the Variables

Variable	Mean	Standard Deviation
General Science (GS)	53.99	7.09
Arithmetic Reasoning (AR)	55.19	6.53
Word Knowledge (WK)	54.53	4.64
Paragraph Comprehension (PC)	54.68	4.68
Numerical Operations (NO)	54.75	6.49
Coding Speed (CS)	54.37	6.77
Auto & Shop Information (AS)	53.40	8.48
Mathematics Knowledge (MK)	55.38	7.76
Mechanical Comprehension (MC)	54.97	7.84
Electronics Information (EI)	53.02	8.41
Two Hand Coordination Horizontal Error (THH)	6963.81	3754.28
Two Hand Coordination Vertical Error (THV)	7846.74	3867.61
Complex Coordination Horizontal Error (CCH)	25914.74	18503.59
Complex Coordination Vertical Error (CCV)	25212.14	20383.90
Complex Coordination Rudder (CCR)	17673.07	17245.62
Time Sharing Response Time (TSR)	1184.18	316.04
Time Sharing Difficulty Without Digit Cancellation (TSS)	252.08	83.96
Time Sharing Difficulty With Digit Cancellation (TSD)	197.62	47.24

Note. The abbreviations presented after the variable are used in all tables.

Contrary to expectations from the literature, there were moderate correlations observed between psychomotor and cognitive tests scores. The average corrected-for-range-restriction correlation between cognitive test scores and psychomotor test scores was .30. It should be noted that these correlations have not been corrected for unreliability and measurement equivalence can not be ascertained from them. The mechanical comprehension test had the highest average correlation with the psychomotor scores at .44; while coding speed had the lowest correlation at .15. Among the psychomotor tests the time sharing test provided scores which were the most and least correlated with the cognitive tests. TSR, time sharing response time, was the most correlated at .41 and TSD, time sharing difficulty with digit cancellation, the least at .22.

Table 3 presents the results of the commonality analyses of the psychomotor tests. Commonality was estimated by regressing each psychomotor test score on the 10 paper-and-pencil cognitive tests. All correlations were statistically significant.

Table 2. Correlations Among the Cognitive and Psychomotor Tests

	GS	AR	WK	PC	NO	CS	AS	MK	MC	EI	THH	THV	CCH	CCV	CCR	TSR	TSS	TSD
GS	**100**	44	59	42	03	06	45	46	50	55	23	22	17	21	18	18	24	22
AR	72	**100**	32	31	25	15	29	59	43	26	19	20	15	22	19	23	18	19
WK	80	71	**100**	53	02	10	30	26	30	40	14	14	06	11	07	12	08	04
PC	69	67	80	**100**	06	14	17	27	20	31	07	06	-01	00	00	14	05	01
NO	52	63	62	61	**100**	55	-13	30	-03	-07	04	01	05	06	12	12	-04	01
CS	45	52	55	56	70	**100**	-08	14	-04	-05	-07	-08	01	02	09	09	-03	-01
AS	64	53	53	42	31	23	**100**	09	60	62	27	22	31	36	24	22	34	29
MK	70	83	67	64	62	52	42	**100**	37	27	17	17	13	16	17	17	22	19
MC	70	68	59	52	41	34	74	60	**100**	59	37	35	34	40	31	28	44	40
EI	76	66	68	57	42	34	75	59	74	**100**	28	26	30	31	21	19	30	24
THH	38	36	31	25	25	11	40	31	47	41	**100**	92	38	43	33	23	51	48
THV	35	34	29	22	21	09	35	30	44	38	93	**100**	41	44	38	21	50	48
CCH	25	26	17	10	18	11	38	22	39	35	42	44	**100**	64	57	31	51	50
CCV	35	37	28	17	25	17	47	31	49	42	49	49	67	**100**	57	25	48	48
CCR	31	33	24	17	27	22	36	29	41	34	39	43	60	61	**100**	27	38	38
TSR	42	46	40	41	39	33	39	41	46	41	33	30	35	34	35	**100**	37	44
TSS	29	27	17	14	10	09	40	27	47	35	54	53	53	51	42	41	**100**	85
TSD	25	24	12	09	11	08	35	22	42	28	50	50	52	51	41	46	86	**100**

Note. Decimal points omitted. Entries above the diagonal are observed data, those below have been corrected-for-range-restriction.

Table 3. **Correlations of Psychomotor Tests with Cognitive Tests and *g***

Test	R	R_c	rg	rg_c	rg_{fc}
THH	.397	.500	.306	.531	.783
THV	.372	.467	.280	.494	.700
CCH	.402	.458	.284	.500	.745
CCV	.456	.546	.344	.581	.959
CCR	.363	.532	.288	.506	.834
TSR	.333	.532	.301	.525	.601
TSS	.467	.507	.331	.565	.642
TSD	.429	.470	.296	.517	.607

Note. R is the multiple correlation between the psychomotor test and the cognitive tests, R_c is R corrected-for-range-restriction. rg is the correlation of g and the psychomotor test, rg_c is rg corrected-for-range-restriction, and rg_{fc} is rg corrected for both range restriction and unreliability. All correlations significant at $p < .01$ Type I error rate.

The average multiple correlation of psychomotor test scores predicted by paper-and-pencil test scores was .40. Corrected-for-range-restriction the average multiple correlation (R_c) was .50. Table 3 also shows the correlation of the psychomotor test scores and psychometric g. The average observed correlation (rg) was .30 and all correlations were statistically significant. These correlations were corrected-for-range-restriction (rg_c) and for the unreliability of both the psychomotor tests and the estimates of g. The average fully corrected correlation (rg_{fc}) was .73.[1]

An unrotated principal components analysis was conducted on the matrix of cognitive and psychomotor test correlations corrected-for-range-restriction. The results (Table 4) revealed a substantial first component, about 46% of the variance, with all the cognitive tests loaded positively, indicating that the component was probably an estimate of psychometric g. Additionally, each of the psychomotor tests had positive loadings on the first component.

To determine if the first unrotated principal component were still a measure of g with the inclusion of the psychomotor tests, a comparison of the g loadings of the ASVAB tests was made with loadings from a previous analysis using only the ASVAB. If the loadings changed proportionally, that is only with regard to scale, then the factor estimated is, in practice, the same. In general the loadings of the ASVAB were proportional in both data sets with an approximately -.10 scaling factor. The coefficient

[1]A meta analysis of these correlations could have disclosed if their variance was the result of real differences or the result of sampling. A meta analysis was not performed because no acceptable estimate of the expected sampling variance of the correlations in Table 3 could be made as the same sample was used for all correlations (Hunter & Schmidt, 1990).

of congruence between the two sets of loadings was .99, indicative of great similarity and further substantiation that the first principal component was *g*.

Table 4. Loadings of the Tests on Psychometric *g* Based on First Unrotated Principal Component

Test	*g* loading	Corrected *g* loading
GS	.8517	.9522
AR	.8467	.9077
WK	.8192	.8782
PC	.7371	.9005
NO	.6759	.7966
CS	.5697	.6493
AS	.7276	.8035
MK	.7886	.8605
MC	.8273	.9428
EI	.8255	.9797
THH	.5854	.8631
THV	.5596	.7615
CCH	.4778	.6826
CCV	.5843	.9239
CCR	.5258	.8313
TSR	.6187	.7734
TSS	.5720	.6241
TSD	.5341	.6009

Note. The loadings in column 2 have been corrected for the unreliability of the tests

The *g* loadings for the psychomotor tests were lower than those for the cognitive tests. These lower loadings were due in part to the lower reliability of the psychomotor tests. When the loadings were divided by the square roots of the reliabilities of the measures, they became more alike. The range of *g* loadings was similar for both types of tests. The average corrected *g* loading was .87 for the cognitive tests and .76 for the psychomotor tests.

The electronics information test showed the highest *g* loading, .97, among the paper-and-pencil cognitive tests, and complex coordination vertical tracking distance error showed the highest *g* loading, .92, among the psychomotor tests. TSD, the time sharing difficulty with digit cancellation test, showed the lowest *g* loading, .60, and coding speed, a paper-and-pencil cognitive test, the second lowest at .64. That the two most highly *g* loaded tests were not verbal and quantitative measures as typically believed is an example of Spearman's "indifference of the indicator" (Spearman, 1927). Jensen (1992) discusses the *g* loading of tests and the indifference of the indicator in detail. It is inappropriate to judge what factors a test measures by its appearance.

The three confirmatory factor analytic models with 3, 7, and 9 factors were estimated. The 3 factor model showed a poor fit to the data as demonstrated by its failure to reach convergence. The factors of *g*, paper-and-pencil tests, and computer tests were inadequate. The 7 and 9 factor models fit equally well as indicated by the Bentler-Bonett non-normed fit indexes both of which were greater than .99. The χ^2/I2 indexes were .99 for both the 7 and 9 factor models showing that each model was a better fit than the null or independence model and that each was a good fit to the data. The TLI incremental fit index was .046 showing no better fit for the 9 factor model than for the 7 factor model. Additionally, the standardized residuals for the 7 and 9 factor models were quite similar. On the basis of the fit statistics, the similarity of the standardized residuals, and parsimony, the 7 factor model was selected as appropriate. The model with 4 lower order paper-and-pencil factors fit no better than the model with only 2 paper-and-pencil factors. Figure 1 shows the structure of the model and Table 5 shows the factor loadings. All factors are orthogonal and residualized.

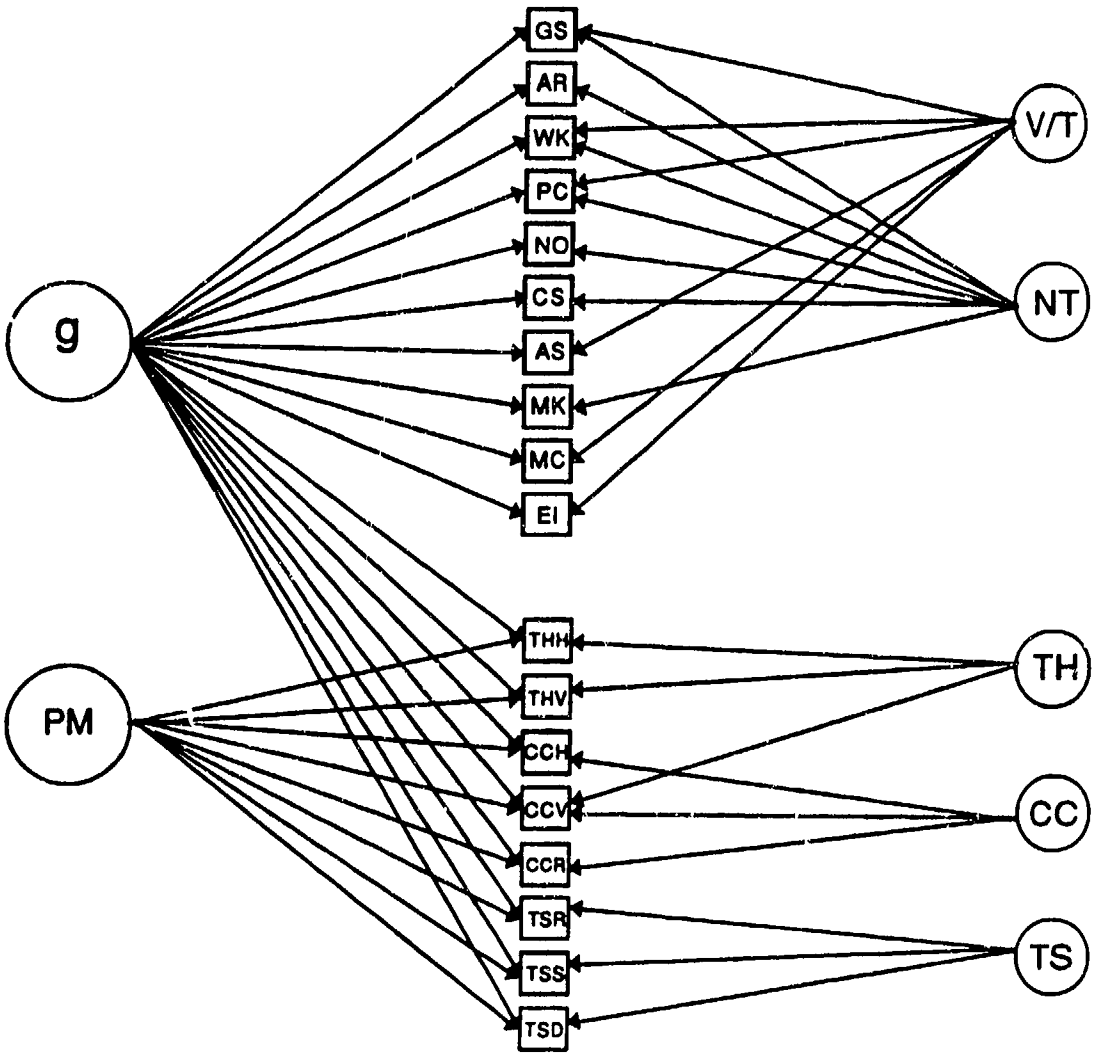

Figure 1. Structural Model of Test Performance from Confirmatory Factory Analysis.

Table 5. Loadings of the Confirmatory Factor Analysis

	Factor						
Test	g	PM	NT	V/T	TH	CC	TS
GS	.771		.175	.432			
AR	.897		.219				
WK	.690		.379	.498			
PC	.643		.409	.363			
NO	.570		.634				
CS	.443		.658				
AS	.768			.371			
MK	.848		.264				
MC	.864			.188			
EI	.796			.391			
THH	.499	.459			.696		
THV	.465	.507			.671		
CCH	.420	.505				.462	
CCV	.531	.399			.053	.477	
CCR	.455	.368				.440	
TSR	.550	.220					.281
TSS	.468	.663					.389
TSD	.416	.648					.562

Note. Factors: g is psychometric *g*, PM is higher order psychomotor, NT is non-technical, V/T is verbal technical, TH is two hand psychomotor, CC is complex coordination psychomotor, and TS is time sharing. g and PM are higher order factors and the others are residualized lower order factors. All are orthogonal. Loading presented are those estimated in the structural equations, all others are essentially zero.

The factors were psychometric *g*, a higher order general psychomotor factor (PM), verbal-technical factor (V/T), a non-technical general knowledge factor (NT), a two hand coordination factor (TH), a complex coordination factor (CC) and a time sharing factor (TS). The proportions of the variance attributable to the higher order factors were, 57% for *g* and 9% for general psychomotor. Among the residualized lower order factors non-technical accounted for 10%; verbal-technical 8%; two hand coordination, 7%, complex coordination, 5%, and time sharing, 4%.

DISCUSSION

The finding that psychomotor tests and cognitive tests had correlated scores was contrary to the expectancy derived from the literature, especially the work of Fleishman and associates (Fleishman, 1953, 1964, 1966, 1972; Fleishman & Quaintance, 1984). It is easy to understand how cognitive tests and psychomotor tests could be seen to be relatively independent by inspection of the uncorrected correlations in Table 2. However, the corrected correlations showed that the two types of measures were not

independent. The correlations between cognitive and psychomotor tests may be due to the requirement to reason (the foundation of *g*) while taking the tests. That mechanical comprehension showed the highest average correlation with psychomotor tests and coding speed the lowest, was probably due to the disparity in reasoning required by these two paper-and-pencil cognitive tests. Mechanical comprehension requires complex reasoning, as do the psychomotor tests, while coding speed requires virtually no reasoning. Further, it is interesting to note that the arithmetic reasoning test, which the confirmatory factor analysis showed as having a greater *g* saturation than mechanical comprehension, did not correlate, on average, as strongly with the psychomotor tests. This may be due to the greater specific saturation of arithmetic reasoning by the NT factor. Mechanical comprehension was nearly as *g* saturated but was less saturated by a specific factor, V/T. Time sharing test scores TSR and TSD were the most, .53, and least, .47, predictable by the cognitive test battery. Time sharing reaction time and task difficulty with digit canceling were collected simultaneously and the differences in predictability between them may be due to the requirement to split attentional resources between the primary (task difficulty) and secondary (reaction time) task. The time sharing test without the secondary task (TSS) showed an average correlation corrected-for-range-restriction with the paper-and-pencil test scores of .50, greater than when attention was shared with a secondary task.

The finding that psychomotor tests were *g* loaded was generally unexpected. Given the importance and use of *g* measures in selection (Hunter & Hunter, 1984; Thorndike, 1985; Ree & Earles, 1991b; Ree, Earles, & Teachout, 1992) the implication of this finding is that the incremental predictive validity of psychomotor tests will be a function of the unique variance of psychomotor skills in the criterion. This is because the *g* component of psychomotor tests is not independent of the *g* component of paper-and-pencil tests and other predictors (Hunter & Hunter, 1984). Ree and Earles (1992) have discussed the efficacy of psychomotor tests for the prediction of pilot performance.

Contrary to Cronbach's (1970) assertion, the confirmatory factor analysis showed a general (higher order) psychomotor factor. As in cognitive tests, this factor might be the source of validity, as might the three lower order psychomotor factors or some combination of higher and lower order factors. The validity of these factors could be important in the development of alternate psychomotor test forms. If the main predictive portion of the psychomotor tests were the higher order factor (PM), then most aggregations of psychomotor tests could be expected to provide a measure of PM. This would facilitate building alternate forms. If however, the validity were a consequence of only the specific psychomotor factors as implied by Fleishman and Quaintance (1984) then the development of alternate forms would be more difficult as new measures of the specific lower order factors were sought. The validity of the general and specific psychomotor factors must be studied in the same manner (Ree & Earles, 1991b) as the validity of paper-and-pencil test factors to understand their role in prediction.

In contrast to common belief, it was found that both the paper-and-pencil and psychomotor tests measured the *g* factor and that there was a higher psychomotor factor. This information is important for future test development and serves as a guide for both research and development and operational testing.

REFERENCES

Bentler, P. M. (1989). *EQS Structural equations program manual.* Los Angeles: CA, BMDP Statistical Software.

Birnbaum, Z, W., Paulson, E., & Andrews, F. C. (1950). On the effects of selection performed on some coordinates of a multi-dimensional population. *Psychometrika, 15,* 191-204.

Burt, C. (1948). The factorial study of temperamental traits. *British Journal of Psychology (Statistical Section), 1,* 3-26.

Carretta, T. R. (1987). *Basic Attributes Test (BAT) system: The development of an automated test battery for pilot selection* (AFHRL-TR-87-9, AD-A186 649). Brooks AFB, TX: Air Force Human Resources Laboratory, Manpower and Personnel Division.

Carretta, T. R. (1989). USAF pilot selection and classification systems. *Aviation Space and Environmental Medicine, 60,* 46-49.

Carretta, T. R. (1990). Cross validation of experimental USAF pilot training performance models. *Military Psychology, 2,* 257-264.

Carretta, T. R. (1992). Short-term retest reliability of an experimental U. S. Air Force pilot candidate selection test battery. *The International Journal of Aviation Psychology, 3,* 161-173.

Carroll, J. (1991). No demonstration that *g* is not unitary, but there's more to the story: Comment on Kranzler and Jensen. *Intelligence, 15,* 423-436.

Cronbach, L. J. (1970). *Essentials of Psychological Testing* (3rd ed.). New York: Harper & Row.

Earles, J. A., & Ree, M. J. (1992a). *Air Force Officer Qualifying Test (AFOQT): Estimating the general ability component* (AL-TP-1991-0039, AD-A245 078). Brooks, AFB, TX: Armstrong Laboratory, Human Resources Directorate, Manpower and Personnel Research Division.

Earles, J. A., & Ree, M. J. (1992b). The predictive validity of the ASVAB for training grades. *Educational and Psychological Measurement, 52,* 721-725.

Fleishman, E. A. (1953). Testing for psychomotor abilities by means of apparatus tests. *Psychological Bulletin, 50,* 241-262.

Fleishman, E. A. (1964). *The structure and measurement of physical fitness.* Englewood Cliffs, NJ: Prentice-Hall.

Fleishman, E. A. (1966). Human abilities and the acquisition of skill. In E. A. Bilodeau (Ed.), *Acquisition of skill.* New York: Academic Press.

Fleishman, E. A. (1972). On the relation between abilities, learning, and human performance. *American Psychologist, 27,* 1017-1032.

Fleishman, E., & Quaintance, M. K. (1984). *Taxonomies of human performance: The description of human tasks.* Orlando, FL: Academic Press.

Hotelling, H. H. (1933a). Analysis of a complex of statistical variables with principal components. *Journal of Educational Psychology, 24,* 417-441.

Hotelling, H. H. (1933b). Analysis of a complex of statistical variables with principal components (continued). *Journal of Educational Psychology, 24,* 498-520.

Hunter, J. E. (1980). *The dimensionality of the General Aptitude Test Battery (GATB) and the dominance of general factors over specific factors in the prediction of job performance.* Washington DC: U. S. Employment Service, U. S. Department of Labor.

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternate predictors of job performance. *Psychological Bulletin, 96,* 72-98.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis.* Newbury Park, CA: Sage.

Jackson, D. N. (1984). *Multidimensional Aptitude Battery Manual.* Port Huron, MI: Research Psychologists Press.

Jackson, D. E., & Ree, M. J. (1990). *Tool for studying the effects of range restriction in correlation coefficient estimation* (AFHRL-TP-90-6, AD-A224 035). Brooks AFB TX: Manpower and Personnel Division.

Jensen, A. R. (1980). *Bias in mental testing.* New York: The Free Press.

Jensen, A. R. (1987). The g beyond factor analysis. In J. C. Conoley, J. A. Glover, & R. R. Ronning (Eds.), *The influence of cognitive psychology on testing and measurement.* Hillsdale, NJ: Erlbaum.

Jensen, A. R. (1992). Commentary: Vehicles of *g*. *Psychological Science, 3,* 275-278.

Kass, R. A., Mitchell, K. J., Grafton, F. C., & Wing, H. (1983). Factorial validity of the Armed Services Vocational Aptitude Battery (ASVAB), Forms 8, 9, and 10. *Educational and Psychological Measurement, 43,* 1077-1087.

Kranzler, J. H., & Jensen, A. R. (1991). The nature of psychometric *g*; Unitary process or a number of independent processes? *Intelligence, 15,* 397-422.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence, 14*, 389-433.

Lawley, D. N. (1943). A note on Karl Pearson's selection formulas. *Proceedings of the Royal Society of Edinburgh.* Section A, 62, Part I, 28-30.

Maier, M. H., & Sims, Sims, W. H. (1986). *The ASVAB score scales: 1980 and World War II.* CNR 116, Alexandria, VA: Center for Naval Analyses.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin, 103*, 391-410.

Neubauer, A. C. (1991). Intelligence and RT: A modified Hick paradigm and a new RT paradigm. *Intelligence, 15*, 175-192.

Raven, J. C. (1966). *Advanced Progressive Matrices.* New York: Psychological Corporation.

Ree, M. J., & Earles, J. A. (1991a). The stability of *g* across different methods of estimation. *Intelligence, 15*, 271-278.

Ree, M. J., & Earles, J. A. (1991b). Predicting training success: Not much more than *g*. *Personnel Psychology, 44*, 321-332.

Ree, M. J., & Earles, J. A. (1992). *The ubiquitous predictiveness of g.* An invited paper for the Army Research Institute sponsored symposium on selection and classification. Washington, DC.

Ree, M. J., Earles, J. A., & Teachout, M. S. (1992). *General cognitive ability predicts job performance* (AL-TP-1991-0057, AD-A245 099). Brooks AFB, TX: Armstrong Laboratory, Human Resources Directorate, Manpower and Personnel Research Division.

Ree, M. J., Mullins, C. J., Mathews, J. J., & Massey, R. H. (1982). *Armed Services Vocational Atitude Battery: Item and factor analysis of Forms 8, 9, and 10* (AFHRL-TR-81-55, AD-A111 430). Brooks, AFB, TX: Air Force Human Resources Laboratory, Manpower and Personnel Division

Ree, M. J., & Wegner, T. G. (1990). Correcting differences in answer sheets for the 1980 Armed Services Vocational Aptitude Battery reference population. *Military Psychology, 2*, 157-169.

Skinner, J., & Ree, M. J. (1987). *Air Force Officer Qualifying Test (AFOQT): Item and factor analysis of Form 0* (AFHRL-TR-86-68, AD-A184 975). Brooks AFB TX: Air Force Human Resources Laboratory, Manpower and Personnel Division.

Spearman, C. (1927). *The abilities of man: Their nature and measurement.* London: Macmillan.

Thorndike, R. L. (1985). The central role of general ability in prediction. *Multivariate Behavior Research, 20,* 241-254.

Thorndike, R. L., & Hagen, E. (1959). *Ten thousand careers.* New York: Wiley.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38,* 1-10.

APPENDIX A

Notes on the Correction for Range Restriction and Its Application to the Issue at Hand

The correlation coefficient is a measure of association and was developed in its current form by Karl Pearson. He built upon the work of Sir Francis Galton who conceived of the idea and on the mathematical formulations of the product-moment of the Austrian mathematician, Bravis. Pearson recognized early on that the sample correlation was a downward biased estimator of the population parameter. Perhaps more importantly he observed that computation of the correlation coefficient in samples which were censured in such a way as to reduce variance in one or both of the variables would seriously downwardly bias the population estimates. He set out a theorem making the assumptions of normality, linearity of form and equality of error variance to show how the correlation coefficients could be adjusted or corrected for the downward bias caused by the censuring. Two of these assumptions are those of linear regression (linearity and equal error variance). This technique has become popularly known as the correction for range restriction. Although Pearson worked out the specific case for two variables at the turn of the century, it was not until 1943 that the English mathematician, Lawley worked out the general case which allows for correction of a number of variables simultaneously, the multivariate correction. Lawley found that the assumption of normality was not needed and that only the two assumptions of linear regression were required. If these two assumptions are not met and the correction can not be made, neither can a linear regression be computed nor can a Pearson product-moment correlation be computed. Assuming that a correlation coefficient can be computed is identical to assuming that the conditions for the correction have been met. That is, if you meet all the assumptions for the correlation you meet all the assumptions for the correction for range restriction.

A question that has arisen in discussions is how does the multivariate correction work? This is especially perplexing in cases where no population information is available on some of the variables. Lawley (1943) provides a proven theorem. Most frequently matrix algebra notation is used to demonstrate the proof although Lawley derived it through the mathematics of moment generation.

The following is taken from Jackson and Ree (1990) and is based on Birnbaum, Paulson, and Andrews (1950).

Assumption 1: (Linearity) For each j the true regression of Y_j on X is linear.

Assumption 2: (Equality of Error Variance). The conditional variance-covariance matrix of Y given X does not depend on X.

Theorem: Under assumptions 1 and 2

$$W_{p,n-p} = W_{p,p} V_{p,p-1} V_{p,n-p} \text{ and}$$

$$W_{n-p,n-p} = V_{n-p,n-p} - V_{n-p,p} (V_{p,p-1} - V_{p,p-1} W_{p,p} V_{p,p-1}) V_{p,n-p}$$

where $W_{p,n-p}$ is the variance-covariance matrix of variables in the unrestricted sample for which no estimates are yet available. $W_{p,p}$ is the variance-covariance matrix of variables in the unrestricted sample for which estimates are available. $V_{p,p}$ is the variance-covariance matrix of variables in the restricted sample and includes the same

variables as $W_{p,p}$. $V_{p,n-p}$ is a variance-covariance matrix in the restricted sample and includes the same variables as $W_{p,n-p}$ but from the restricted sample. $W_{n-p,n-p}$ is the estimated (corrected) variance-covariance matrix of variables for which there were no unrestricted sample estimates. $V_{n-p,n-p}$ is the analogous matrix in the restricted sample for which estimates were available.

This can be more easily shown as block diagrams with the first being for the estimates in the restricted sample and the second for the unrestricted sample. Each represents a variance-covariance matrix (or a standardized variance-covariance matrix, a correlation matrix).

Block 1 Variance-covariance matrix from a restricted sample.

$$\begin{matrix} V_{p,p} & V_{n-p,p} \\ V_{p,n-p} & V_{n-p,n-p} \end{matrix}$$

Block 2 Variance-covariance matrix from an unrestricted sample.

$$\begin{matrix} W_{p,p} & W_{n-p,p} \\ W_{p,n-p} & W_{n-p,n-p} \end{matrix}$$

All the variances and covariances are known for the Vs. Only $W_{p,p}$ is known in Block 2 but given the theorem and the equations, $W_{p,n-p}$ and $W_{n-p,n-p}$ can be estimated. $W_{n-p,p}$ is identical to $W_{p,n-p}$ with rows and columns transposed. Said differently, knowledge of some of the variables in the unrestricted sample ($W_{p,p}$) and all in the restricted sample (all the Vs) coupled with the two equations of the theorem allows all the variances (the remaining Ws) to be estimated in the unrestricted sample. This explains how negative correlations in the restricted sample can be positive in the unrestricted sample. It follows from the laws and consequences of matrix multiplication. Consider a negative correlation in $V_{p,p}$ but positive in $W_{p,p}$. The theorem above shows that $W_{p,p}$ will multiply the Vs and correct the Vs This will also occur in the second equation. A sample matrix using the same notation is provided as Table 1. Finally, it is important to note that the multivariate correction is <u>not</u> equivalent to a series of bivariate corrections.

By way of concrete example assume that variables a and b are aptitude tests and c and d are psychomotor tests. In matrix V there is information about all the correlations in the restricted sample. These we believe to be severely downwardly biased estimators and we would be better off with the Ws. The only information we start with in matrix W is for the two aptitude tests. This information is found in $W_{p,p}$. Applying the two equation from the theorem we can solve for the missing parts of W-- $W_{p,n-p}$ and $W_{n-p,n-p}$. $W_{n-p,p}$ is solved for when we solve for $W_{p,n-p}$ because it is the transpose of $W_{p,n-p}$. Applying the two equations gives all unknown parts of the matrix. The interested reader is directed to Birnbaum et al. for a numerical example.

Table A-1. Variance-Covariance (or Correlation) Matrices Using Birnbaum et al. Notation

The restricted matrix.

$$V = \begin{array}{cccc} p,p & n\text{-}p,p & & \\ r_{aa} & r_{ab} & r_{ac} & r_{ad} \\ r_{ba} & r_{bb} & r_{bc} & r_{bd} \\ p,n\text{-}p & & n\text{-}p,n\text{-}p & \\ r_{ca} & r_{cb} & r_{cc} & r_{cd} \\ r_{da} & r_{db} & r_{dc} & r_{dd} \end{array}$$

The unrestricted matrix.

$$W = \begin{array}{cccc} p,p & & n\text{-}p,p & \\ r_{aa} & r_{ab} & r_{ac} & r_{ad} \\ r_{ba} & r_{bb} & r_{bc} & r_{bd} \\ p,n\text{-}p & & n\text{-}p,n\text{-}p & \\ r_{ca} & r_{cb} & r_{cc} & r_{cd} \\ r_{da} & r_{db} & r_{dc} & r_{dd} \end{array}$$

ARMSTRONG LABORATORY
HUMAN RESOURCES DIRECTORATE
MANPOWER AND PERSONNEL RESEARCH DIVISION
Brooks Air Force Base, TX 78235-5000

ERRATA

Ree, Malcolm James, & Carretta, Thomas R. (1992). The Correlation of Cognitive and Psychomotor Tests (AL-TP-1992-0037). Brooks AFB, TX: Manpower and Personnel Research Division, Armstrong Laboratory.

Corrected pages 18 and 20 are attached to replace the ones printed in the original technical paper.

Marion E. Green
MARION E. GREEN
Chief, Editing Services Branch (AL/DOKE)

The correlation coefficient is a measure of association and was developed in its current form by Karl Pearson. He built upon the work of Sir Francis Galton who conceived of the idea and on the mathematical formulations of the product-moment of the Austrian mathematician, Bravis. Pearson recognized early on that the sample correlation was a downward biased estimator of the population parameter. Perhaps more importantly he observed that computation of the correlation coefficient in samples which were censured in such a way as to reduce variance in one or both of the variables would seriously downwardly bias the population estimates. He set out a theorem making the assumptions of normality, linearity of form and equality of error variance to show how the correlation coefficients could be adjusted or corrected for the downward bias caused by the censuring. Two of these assumptions are those of linear regression (linearity and equal error variance). This technique has become popularly known as the correction for range restriction. Although Pearson worked out the specific case for two variables at the turn of the century, it was not until 1943 that the English mathematician, Lawley worked out the general case which allows for correction of a number of variables simultaneously, the multivariate correction. Lawley found that the assumption of normality was not needed and that only the two assumptions of linear regression were required. If these two assumptions are not met and the correction can not be made, neither can a linear regression be computed nor can a Pearson product-moment correlation be computed. Assuming that a correlation coefficient can be computed is identical to assuming that the conditions for the correction have been met. That is, if you meet all the assumptions for the correlation you meet all the assumptions for the correction for range restriction.

A question that has arisen in discussions is how does the multivariate correction work? This is especially perplexing in cases where no population information is available on some of the variables. Lawley (1943) provides a proven theorem. Most frequently matrix algebra notation is used to demonstrate the proof although Lawley derived it through the mathematics of moment generation.

The following is taken from Jackson and Ree (1990) and is based on Birnbaum, Paulson, and Andrews (1950).

Assumption 1: (Linearity) For each j the true regression of Y_j on X is linear.

Assumption 2: (Equality of Error Variance). The conditional variance-covariance matrix of Y given X does not depend on X.

Theorem: Under assumptions 1 and 2

$W_{p,n-p} = W_{p,p} V_{p,p}^{-1} V_{p,n-p}$ and

$W_{n-p,n-p} = V_{n-p,n-p} - V_{n-p,p} (V_{p,p}^{-1} - V_{p,p}^{-1} W_{p,p} V_{p,p}^{-1}) V_{p,n-p}$

where $W_{p,n-p}$ is the variance-covariance matrix of variables in the unrestricted sample for which no estimates are yet available. $W_{p,p}$ is the variance-covariance matrix of variables in the unrestricted sample for which estimates are available. $V_{p,p}$ is the variance-covariance matrix of variables in the restricted sample and includes the same

Table A-1. Variance-Covariance (or Correlation) Matrices Using Birnbaum et al. Notation

The restricted matrix.

	p,p		n-p,p	
	r_{aa}	r_{ab}	r_{ac}	r_{ad}
V=	r_{ba}	r_{bb}	r_{bc}	r_{bd}
	p,n-p		n-p,n-p	
	r_{ca}	r_{cb}	r_{cc}	r_{cd}
	r_{da}	r_{db}	r_{dc}	r_{dd}

The unrestricted matrix.

	p,p		n-p,p	
	r_{aa}	r_{ab}	r_{ac}	r_{ad}
W=	r_{ba}	r_{bb}	r_{bc}	r_{bd}
	p,n-p		n-p,n-p	
	r_{ca}	r_{cb}	r_{cc}	r_{cd}
	r_{da}	r_{db}	r_{dc}	r_{dd}

Preceding Page Blank